

# Explicit computations for some Markov modulated counting processes

Michel Mandjes      Peter Spreij

April 12, 2016

## Abstract

In this paper we present elementary computations for some *Markov modulated* counting processes, also called counting processes with *regime switching*. Regime switching has become an increasingly popular concept in many branches of science. In finance, for instance, one could identify the background process with the ‘state of the economy’, to which asset prices react, or as an identification of the varying default rate of an obligor. The key feature of the counting processes in this paper is that their intensity processes are functions of a finite state Markov chain. This kind of processes can be used to model default events of some companies.

Many quantities of interest in this paper, like conditional characteristic functions, can all be derived from conditional probabilities, which can, in principle, be *analytically* computed. We will also study limit results for models with rapid switching, which occur when inflating the intensity matrix of the Markov chain by a factor tending to infinity. The paper is largely expository in nature, with a didactic flavor.

**Keywords:** Counting process, Markov chain, Markov modulated process, Regime switching.

**AMS subject classification:** 60G44, 60G55, 60J27.

## 1 Introduction

In this paper we present some elementary computations concerning some *Markov modulated* (MM) counting processes, denoted  $N$ , also called counting processes with *regime switching*. Such processes fall into the class of *hybrid models* [29] and are in fact Hidden Markov processes [10]. Although in the present paper we restrict ourselves to certain counting processes, it is worth mentioning that owing to its various attractive features, regime switching has become an increasingly popular concept in many branches of science. In a broad spectrum of application domains it offers a natural framework for modeling situations in which the stochastic process under study reacts to an autonomously evolving environment. In finance, for instance, one could identify the background process with the ‘state of the economy’, to which asset prices react, or as an identification of the varying default rate of an obligor. In operations research, in particular in wireless networks, the concept can be used to model the channel conditions that vary in time, and to which users react. In the literature in the latter field there is a sizeable body of work on Markov-modulated queues, see e.g. [2, Ch. XI] and

[27], while Markov modulation has been intensively used in insurance and risk theory as well [3]. In the economics literature, the use of regime switching dates back to at least the late 1980s [16]. Various specific models have been considered since then, see for instance [1, 11, 12]. For other direct applications of models with regime switching in finance (hedging of claims, interest rate models, credit risk, application to pension funds) we refer to [8, 22, 23, 30, 31] for recent results.

The key feature of the counting processes, commonly denoted  $N$ , in this paper is that their intensity processes are of the form  $\lambda_t = \lambda(X_t, N_t)$ , where  $X$  is a finite state Markov chain whose jumps with probability one never coincide with the jumps of the counting process. For mathematical convenience we assume without loss of generality that  $X$  takes its values in the set of  $d$ -dimensional basis vectors.

This kind of processes can be used to model default events of some companies. We restrict our treatment to models where the intensity is of a special form, leading to the MM one point process which can be used to model the default event of a single company, its extension to the situation of defaults of various companies and an MM Poisson process, which can be used to model defaults for a large pool of obligors whose individual intensities of default are all the same and small.

The intensities  $\lambda_t = \lambda(X_t, N_t)$  that we use will be affine in  $X_t$ , i.e.  $\lambda_t = \lambda^\top X_t f(N_t)$  for some  $\lambda \in \mathbb{R}^d$  and some function  $f$ . It is possible to show that the joint process  $(X, N)$  is Markov, in fact it is an affine process after a state transformation. This means that for many quantities of interest, like conditional characteristic functions, one can in principle use the full technical apparatus that has become available for affine process, see [9]. However, as these quantities can all be derived from conditional probabilities (our processes are finite, or at most countably, valued), using these techniques is like making a detour since the conditional probabilities can be derived by more straightforward methods. Moreover these conditional probabilities give a *direct* insight into the probabilistic structure of the process and can in principle be *analytically* computed. Therefore, we circumvent the theory of affine processes and focus on direct computation of all conditional probabilities of interest.

We will also study limit results for models with rapid switching, which occur when inflating the intensity matrix of the Markov chain by a factor tending to infinity. Rapid switching between (macro) economic states is unrealistic, but in models for the profit and loss of trading positions, especially in high frequency trading, rapid switching may take place, see [15]. We will see that the limit processes have intensities that are expectations under the invariant distribution of the chain. This is similar to what happens in the context of Markov modulated Ornstein-Uhlenbeck processes [18], see also [19], whereas comparable results under scaling in the operations research literature can be found in [5] and [6].

The paper is largely expository in nature, with a didactic flavor. We do not claim novelty of all results below. Rather we emphasize the uniform approach that we follow, using martingale methods, that may also lead to alternative proofs of known results, e.g. those concerning transition probabilities by using ‘ $\varepsilon$ -arguments’ as in [27]. The organization of the paper is as follows. In Section 2 we study Markov modulated model for the total number of defaults when there are  $n$  obligors. As a primer, in Section 2.1 we extensively study the Markov modulated model for a single obligor, in particular its distributional properties.

Then we switch to the more general situation of Section 2.2, where our approach is inspired by the easier case of the previous section. All results are basically obtained by exploiting the Markovian nature of the joint process  $(X, N)$ . Section 3 gives a few results for the Markov modulated Poisson process. Conditional probabilities of future values of the counting processes, when only its own past can be observed (and not the underlying Markov chain) can be computed using filtering theory, which is the topic of Section 4. In Section 5 we obtain the limit results for processes where the Markov chain is rapidly switching.

## 2 The MM model for multiple obligors

We assume throughout that a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  is given. Suppose we have  $n$  obligors with default times  $\tau^i$  for obligor  $i$ ,  $i = 0, \dots, n$ . Let  $Y_t^i = \mathbf{1}_{\{\tau^i \leq t\}}$ ,  $t \in [0, \infty)$ . Here we encounter the canonical set-up for the intensity based approach in credit risk modelling, see [13, Chapter 12] or [4, Chapter 6] for further details on probabilistic aspects. We postulate for each  $i \in \{1, \dots, n\}$

$$dY_t^i = \lambda_t(1 - Y_t^i) dt + dm_t^i, \quad (1)$$

for  $\lambda_t$  a nonnegative process to be specified, but which is the same for each obligor  $i$ . Here each  $m^i$  is a martingale w.r.t. to the filtration, call it  $\mathbb{F}^i$ , generated by  $Y^i$  and the process  $\lambda$ . We impose that the  $\tau_i$  are conditionally independent given  $\lambda$ . Hence, simultaneous defaults occur with probability zero, as the  $\tau^i$  have a continuous distribution. By the conditional independence assumption, the  $m^i$  are also martingales w.r.t.  $\mathbb{F} = \bigvee_{i=1}^n \mathbb{F}^i$ . The process  $\lambda$  is assumed to be predictable w.r.t.  $\mathbb{F}$ . In all what follows in this section we take  $N_t = \sum_{i=1}^n Y_t^i$ .

### 2.1 The MM one point process

For a better understanding of what follows, we single out the special case  $n = 1$  and we write  $\tau$  instead of  $\tau^1$ . There is some advantage in starting with a simpler case that allows for more explicit formulas, is more transparent, and that at the same time can serve as a warming up for the more general setting.

#### 2.1.1 The general one point process with intensity

Let us consider the basic case, the random variable  $\tau$  has an exponential distribution with parameter  $\lambda$ , and  $Y_t = \mathbf{1}_{\{\tau \leq t\}}$ ,  $t \in [0, \infty)$ . Then  $Y$  has semimartingale decomposition

$$dY_t = \lambda(1 - Y_t) dt + dm_t, \quad (2)$$

where  $\lambda > 0$  and  $m$  a martingale w.r.t. the filtration generated by the process  $Y$ . As a matter of fact, the distributional property of  $\tau$  is equivalent to the decomposition of  $Y$  in (2). Clearly  $Y_t$  is a Bernoulli random variable, so  $y(t) := \mathbb{E}Y_t = \mathbb{P}(Y_t = 1) = \mathbb{P}(\tau \leq t)$ . Alternatively, taking expectations, we get the ODE

$$\dot{y}(t) = \lambda(1 - y(t)),$$

which is, with  $y(0) = 0$ , indeed solved by

$$y(t) = 1 - \exp(-\lambda t).$$

Let  $\Lambda^\tau$  be the compensator of  $Y$ , then

$$\Lambda_t^\tau = \int_0^t \lambda(1 - Y_s) ds = \int_0^t \lambda \mathbf{1}_{\{s < \tau\}} ds = \int_0^{t \wedge \tau} \lambda ds = \lambda(\tau \wedge t).$$

Note that  $Y$  can be considered as  $N^\tau$ , the at  $\tau$  stopped Poisson process with intensity  $\lambda$ . The compensator  $\Lambda$  of  $N$  stopped at  $\tau$  indeed yields  $\Lambda^\tau$ .

As a first generalization we change the above setup in the sense that we postulate

$$dY_t = \lambda_t(1 - Y_t) dt + dm_t, \quad (3)$$

where  $\lambda$  is a nonnegative locally integrable Borel function, also known as the (time varying) hazard rate. As above one can show that

$$y(t) = 1 - \exp\left(-\int_0^t \lambda_s ds\right).$$

In a next generalization we suppose that  $\lambda$  becomes a random process defined on an auxiliary probability space  $(\Omega', \mathcal{F}', \mathbb{P}')$ . We can look at the product probability space  $(\Omega \times \Omega', \mathcal{F} \otimes \mathcal{F}', \mathbb{P} \otimes \mathbb{P}')$  and redefine in the obvious way  $Y$ ,  $\tau$  and  $\lambda$  on this product space. The filtration we will use consists of the  $\sigma$ -algebras  $\mathcal{F}_t^Y \otimes \mathcal{F}_t^\lambda$ .

It is assumed that  $\lambda$  is predictable and a.s. locally integrable w.r.t. Lebesgue measure. For a given trajectory  $\lambda_t = \lambda_t(\omega')$  we define  $Y$  on  $(\Omega, \mathcal{F}, \mathbb{P})$  as in (3). With  $\mathcal{F}^\lambda$  the  $\sigma$ -algebra generated by the full process  $\lambda$ , we have that

$$\mathbb{E}[Y_t | \mathcal{F}^\lambda] = 1 - \exp\left(-\int_0^t \lambda_s ds\right),$$

and hence

$$y(t) = \mathbb{E}Y_t = 1 - \mathbb{E} \exp\left(-\int_0^t \lambda_s ds\right).$$

Alternatively, one can construct the point process  $Y$  as follows. Let  $(\Omega, \mathcal{F}, \mathbb{Q})$  be a probability space on which is defined a standard Poisson process  $Y$  and *independently* of  $Y$  the nonnegative predictable process  $\lambda$ . Put  $L_t = \mathcal{E}(\mu)_t$ , the Doléans exponential of the  $\mathbb{Q}$ -local martingale  $\mu$  given by  $\mu_t = \int_0^t (\lambda_s \mathbf{1}_{\{Y_{s-}=0\}} - 1) d(Y_s - s)$ . Note that  $L_0 = 1$ . Let  $\tau_k$  be the consecutive jump times of  $Y$ ,  $\tau_0 = 0$ . Note that the differences  $\tau_k - \tau_{k-1}$  have a standard exponential distribution under  $\mathbb{Q}$ . The assertion of the following lemma is a variation on Equation (4.23) in [4].

**Lemma 2.1** *The density process  $L$  allows the following explicit expression,*

$$L_t = (\lambda_{\tau_1})^{Y_t} \exp\left(t - \int_0^{\tau_1 \wedge t} \lambda_s ds\right) \mathbf{1}_{\{Y_t \leq 1\}}.$$

*If  $\lambda$  is a bounded process,  $L$  is a martingale, hence  $\mathbb{E}L_t = L_0 = 1$ .*

**Proof** By construction,  $L$  is a local martingale. For bounded  $\lambda$  we have  $\mathbb{E} \int_0^t L_s^2 ds \leq C \exp(2t)$  for some constant  $C$ , which yields  $L$  a square integrable martingale. The given expression for  $L_t$  can be verified by an elementary, but slightly tedious computation.  $\square$

Under the assumption that  $L$  is a martingale (guaranteed for bounded  $\lambda$ ), by Girsanov's theorem, see [7, Chapter VI, T3 and T4], we can define for every  $T > 0$  a probability  $\mathbb{P}$  on  $(\Omega, \mathcal{F}_T)$  such that

$$m_t := Y_t - t - \langle Y, \mu \rangle_t = Y_t - \int_0^t \lambda_s \mathbf{1}_{\{Y_{s-}=0\}} ds$$

is a local martingale under  $\mathbb{P}$ . Note that  $\mathbb{P}(Y_T > 1) = \mathbb{E}_{\mathbb{Q}} \mathbf{1}_{\{Y_T > 1\}} L_T = 0$ . Hence, under  $\mathbb{P}$  we have  $\mathbf{1}_{\{Y_s=0\}} = 1 - Y_s$  and the expression for  $m_t$  coincides with (3) for  $t \leq T$ . Note that  $L$  cannot be uniformly integrable, since  $L_\infty = 0$ , which follows from  $L_{\tau_2} = 0$ . Hence it is not automatic that one can define a probability  $\mathbb{P}$  on  $(\Omega, \mathcal{F})$  such that  $m$  is a martingale on  $[0, \infty)$ . Note that the laws under  $\mathbb{P}$  and  $\mathbb{Q}$  of  $\lambda$  are the same.

### 2.1.2 The one point process with MM intensity

In this section we consider (3), where we specify  $\lambda_t$  as a function of a finite state Markov chain  $X_t$ , i.e.  $\lambda_t = \lambda(X_t)$ . We see that, trivial cases excluded, unlike the constant hazard rate  $\lambda$  in (2), we now have a rate that assumes different values according to the states of the Markov chain. We thus have a rate that is subject to *regime switching*, one also says that we have a *Markov modulated* rate. In order to pose a precise mathematical model, we make some conventions. Let  $d$  be the size of the state space of the Markov chain  $X$ . Then w.l.o.g. we may assume that  $X$  takes its values in the set  $\{e_1, \dots, e_d\}$  of  $d$ -dimensional standard basis vectors. This implies that any function of  $X_t$  can be written as a linear map of  $X_t$ , in particular  $\lambda(X_t) = \lambda^\top X_t$ , where on the right hand side  $\lambda$  is a vector in  $\mathbb{R}_+^d$ .

Let  $Q$  be the transition matrix of  $X$ , for which we use the convention that  $Q_{ij}$  for  $i \neq j$  is the intensity of a transition from state  $j$  to state  $i$ . As a consequence the *column sums* of  $Q$  are equal to zero. We then have

$$dX_t = QX_t dt + dM_t^X,$$

where  $M^X$  a martingale with values in  $\mathbb{R}^d$ . We also assume that  $Q$  is irreducible and we denote by  $\pi$  the vector representing the invariant distribution.

Furthermore it will be throughout assumed that  $Y$  and  $X$  have no simultaneous jumps, hence the quadratic variation process  $[X, Y]$  ( $[X, Y]_t = \sum_{s \leq t} \Delta X_s \Delta Y_s$ ) is identically zero.

For the single obligor case, we pose the following model with regime switching,

$$dY_t = \lambda^\top X_t (1 - Y_t) dt + dm_t,$$

where  $\lambda \in \mathbb{R}_+^d$ .

One way of constructing this model is by realizing it on a product space with  $\lambda_t = \lambda^\top X_t$  as in Section 2.1.1. Alternatively, one can realize  $Y$  as standard Poisson process and independently of it,  $X$  as a Markov chain on the auxiliary space under  $\mathbb{Q}$ . By independence, one has  $[X, Y] = 0$  under  $\mathbb{Q}$  and as these brackets remain the same under an absolutely continuous change of measure using the  $\mathbb{Q}$ -martingale  $\mu$  of the previous section, we are then guaranteed to have  $[X, Y] = 0$  under  $\mathbb{P}$  as well. In this case it is possible to have  $\mathbb{P}$  defined on  $(\Omega, \mathcal{F})$  for  $\mathcal{F} = \mathcal{F}_\infty$ , where we use the filtration generated by  $Y$  and  $X$ . As a

side remark we note that  $\mathbb{P}$  will not be absolutely continuous w.r.t.  $\mathbb{Q}$  on  $\mathcal{F}_\infty$ .

In all what follows in this paper we adopt the following *Convention*: we will use the generic notation  $M$  for a martingale, possibly even of varying dimensions, whose precise form is not important.

An important role will be played by the matrices  $Q_{k\lambda} := Q - k \text{diag}(\lambda)$  for  $k \geq 0$ . Here  $\text{diag}(\lambda)$  is the diagonal matrix with  $ii$ -element equal to  $\lambda_i$ . Here is a, possibly known, stability result for the matrix  $Q_\lambda$  (we take  $k = 1$ , but a similar result is obviously true for all positive  $k$ ).

**Lemma 2.2** *Let  $\lambda_i > 0$  for all  $i$ . Then the matrix  $Q_\lambda$  is invertible and  $\exp(Q_\lambda t) \rightarrow 0$  for  $t \rightarrow \infty$ .*

**Proof** That  $Q_\lambda$  is invertible, can be seen as follows. Write

$$Q_\lambda = -(I - Q \text{diag}(\lambda)^{-1}) \text{diag}(\lambda)$$

and note that  $Q \text{diag}(\lambda)^{-1}$  is also the intensity matrix of a Markov chain, as its off-diagonal elements are positive and  $\mathbf{1}^\top Q \text{diag}(\lambda)^{-1} = 0$ . Therefore  $I - Q \text{diag}(\lambda)^{-1}$  is invertible, and so is  $Q_\lambda$ .

In proving the limit result, we give a probabilistic argument.<sup>1</sup> Consider the augmented matrix

$$Q_\lambda^a = \begin{pmatrix} 0 & -\mathbf{1}^\top Q_\lambda \\ 0 & Q_\lambda \end{pmatrix},$$

which is the transition matrix of a Markov chain taking values in  $\{e_0^a, \dots, e_d^a\}$ , labelled as the standard basis vectors of  $\mathbb{R}^{d+1}$ . Clearly, 0 is an absorbing state, and the only one. Hence whatever initial state  $x^a(0)$ , we have that  $\exp(Q_\lambda^a t) x^a(0) \rightarrow e_0^a$  for  $t \rightarrow \infty$ . Computing the exponential and taking  $x^a(0) \neq e_0^a$ , we find

$$\exp(Q_\lambda^a t) x^a = \begin{pmatrix} 1 & \mathbf{1}^\top (I - \exp(Q_\lambda t)) \\ 0 & \exp(Q_\lambda t) \end{pmatrix} x^a(0) = \begin{pmatrix} \mathbf{1}^\top (I - \exp(Q_\lambda t)) x(0) \\ \exp(Q_\lambda t) x(0) \end{pmatrix}$$

Hence  $\exp(Q_\lambda t) \rightarrow 0$ . □

In a next section, see Remark 2.7, we shall see how to compute  $\mathbb{P}(Y_t = 1)$ . It turns out to be the case that

$$\mathbb{P}(Y_t = 1) = 1 - \mathbf{1}^\top \exp(Q_\lambda t) x(0).$$

We conclude in view of Lemma 2.2 that  $\mathbb{P}(Y_t = 1) \rightarrow 1$  for  $t \rightarrow \infty$ . Hence, with probability one, the obligor eventually defaults, as expected.

## 2.2 The MM model for multiple obligors

In Section 2.1.2 we have seen results for default processes in the situation of a single obligor. In the present section we generalize those results, at the cost of considerably more complexity, to the situation of multiple obligors.

---

<sup>1</sup>This argument has been provided by Koen de Turck, University of Ghent.

### 2.2.1 Multiple obligors with time-varying intensity

Recall (1). Let's first look at the constant intensity case,  $\lambda_t = \lambda > 0$ . Then  $N_t = \sum_{i=1}^n Y_t^i$  satisfies

$$dN_t = \lambda(n - N_t) dt + dm_t, \quad (4)$$

where  $m = \sum_{i=1}^n m^i$ . By the independence of the default times,  $m$  is a martingale w.r.t.  $\mathbb{F}$  and  $N_t$  has the  $\text{Bin}(n, 1 - \exp(-\lambda t))$  distribution. Moreover, given  $N_u$ ,  $u \leq s$ ,  $N_t - N_s$  has for  $t > s$  the  $\text{Bin}(n - N_s, 1 - \exp(-\lambda(t - s)))$  distribution. This model has long ago been used in software reliability going back to [21], with various refinements, like in a Bayesian set up the parameters  $n$  and  $\lambda$  being random, see [25, 26] or with time varying but deterministic intensity function  $\lambda(t)$ , see [14].

Next we look at the case of time varying, possibly random,  $\lambda$ . By the assumed conditional independence of the  $\tau^i$  given  $\lambda$  we have, similar to the constant  $\lambda$  case, that  $N_t$ , conditional on the process  $\lambda$ , has a  $\text{Bin}(n, 1 - \exp(-\Lambda_t))$  distribution with  $\Lambda_t = \int_0^t \lambda_s ds$ .

Let  $p^k(t) = \mathbb{P}(N_t = k | \mathcal{F}^\lambda)$ , put

$$p(t) = \begin{pmatrix} p^0(t) \\ \vdots \\ p^n(t) \end{pmatrix}$$

and

$$A = \begin{pmatrix} -n & 0 & \cdots & \cdots & \cdots & 0 \\ n & -(n-1) & 0 & \cdots & \cdots & 0 \\ 0 & n-1 & -(n-2) & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & -1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 & 0 \end{pmatrix}. \quad (5)$$

Then we have for  $p(t)$  the system of differential equations

$$\dot{p}(t) = \lambda_t A p(t),$$

which has solution (here we use that  $\lambda$  is real-valued)

$$p(t) = \exp(\Lambda_t A) e_0,$$

where  $\Lambda_t = \int_0^t \lambda_s ds$  and  $e_0$  is the first standard basis vector of  $\mathbb{R}^{n+1}$ . For the vector whose elements are the unconditional probabilities  $\mathbb{P}(N_t = k)$  one has to take the expectation and it depends on the specification of  $\lambda$  whether this results in analytic expressions. We will see that this happens in case of a Markov modulated rate process.

### 2.2.2 The MM case

We assume to have a finite state Markov process as in Section 2.1.2 and let  $\lambda_t = \lambda^\top X_{t-}$ . For  $N_t$  one now has its submartingale decomposition

$$dN_t = \lambda^\top X_t (n - N_t) dt + dm_t.$$

This is the model of Section 2.1.2 extended to more obligors. The default rate for each obligor has become random ( $\lambda^\top X_t$ ), but is taken the same for all of them.

Let  $\nu_t^k = \mathbf{1}_{\{N_t=k\}}$ ,  $k = 0, \dots, n$ . For notational convenience we set  $\nu_t^{-1} = 0$ . It follows that  $\Delta \nu_t^k = 1$  iff  $N_t$  jumps from  $k-1$  to  $k$  at  $t$ , and  $\Delta \nu_t^k = -1$  iff  $N_t$  jumps from  $k$  to  $k+1$ . This can be summarized by

$$d\nu_t^k = (\nu_{t-}^{k-1} - \nu_{t-}^k) dN_t.$$

In vector form this becomes

$$d\nu_t = (J - I)\nu_{t-} dN_t, \quad (6)$$

where

$$J = \begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ 0 & 1 & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}.$$

Using the dynamics for  $N$ , we get

$$\begin{aligned} d\nu_t^k &= (\nu_{t-}^{k-1} - \nu_{t-}^k)(\lambda^\top X_{t-}(n - N_t) dt + dm_t) \\ &= \lambda^\top X_t((n - k + 1)\nu_t^{k-1} - (n - k)\nu_t^k) dt + dM_t. \end{aligned}$$

Letting  $\nu_t = \begin{pmatrix} \nu_t^0 \\ \vdots \\ \nu_t^n \end{pmatrix}$ , we get from the above display

$$d\nu_t = \lambda^\top X_t A \nu_t dt + dM_t, \quad (7)$$

where  $A$  is as in (5). This equation for  $\nu$  is a main ingredient in the next result.

**Proposition 2.3** *Let  $\zeta_t = \nu_t \otimes X_t$ . The process  $\zeta$  is Markov with transition matrix  $\mathbf{Q}$ , where  $\mathbf{Q} = (A \otimes \text{diag}(\lambda) + I \otimes Q)$ . It follows that  $\mathbb{E}[\zeta_t | \mathcal{F}_s] = \exp(\mathbf{Q}(t - s))\zeta_s$ .*

**Proof** We will use equation (7) together with the dynamics of  $X$ . Using the product rule and the fact that  $N$  and  $X$  do not jump at the same time and summarizing again all martingale terms again as  $M$ , we get (recall the multiplication rule  $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ )

$$\begin{aligned} d(\nu_t \otimes X_t) &= ((A\nu_t \lambda^\top X_t) \otimes X_t + \nu_t \otimes (QX_t)) dt + dM_t \\ &= ((A\nu_t) \otimes (X_t \lambda^\top X_t) + \nu_t \otimes (QX_t)) dt + dM_t \\ &= ((A\nu_t) \otimes (\text{diag}(\lambda)X_t) + I\nu_t \otimes (QX_t)) dt + dM_t \\ &= (A \otimes \text{diag}(\lambda) + I \otimes Q)(\nu_t \otimes X_t) dt + dM_t \\ &= \mathbf{Q}(\nu_t \otimes X_t) dt + dM_t. \end{aligned}$$

Note that  $\zeta_t$  by construction consists of the indicators of the values of the joint process  $(\nu, X)$ . Hence the equation  $d\zeta_t = \mathbf{Q}\zeta_t dt + dM_t$  reveals, cf. Lemma 1.1 in Appendix B of [10], that  $\zeta$  (and hence  $(\nu, X)$ ) is Markov.  $\square$



An explicit computation shows

$$\mathbf{Q} = \begin{pmatrix} Q_{n\lambda} & 0 & \cdots & \cdots & \cdots & 0 \\ n \operatorname{diag}(\lambda) & Q_{(n-1)\lambda} & 0 & \cdots & \cdots & 0 \\ 0 & (n-1) \operatorname{diag}(\lambda) & Q_{(n-2)\lambda} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & Q_\lambda & 0 \\ 0 & \cdots & \cdots & 0 & \operatorname{diag}(\lambda) & Q \end{pmatrix}, \quad (8)$$

where for  $k \in \mathbb{N}$  we have  $Q_{k\lambda} = Q - k \operatorname{diag}(\lambda)$ .

**Remark 2.4** The original dynamic equations for  $X_t$  and  $N_t$  can be retrieved from Proposition 2.3. Realizing the relations  $X_t = (\mathbf{1}^\top \otimes I)\zeta_t$  and  $(\mathbf{1}^\top \otimes I)\mathbf{Q} = \mathbf{1}^\top \otimes Q$ , and  $\mathbf{1}^\top A = 0$ , we obtain from Proposition 2.3

$$\begin{aligned} dX_t &= (\mathbf{1}^\top \otimes I)(\mathbf{Q}(\nu_t \otimes X_t)) dt + dM_t \\ &= (\mathbf{1}^\top \otimes Q)(\nu_t \otimes X_t) dt + dM_t \\ &= QX_t dt + dM_t. \end{aligned}$$

Similarly, we get from  $\nu_t = (I \otimes \mathbf{1}^\top)\zeta_t$ ,

$$\begin{aligned} d\nu_t &= (I \otimes \mathbf{1}^\top)(\mathbf{Q}(\nu_t \otimes X_t)) dt + dM_t \\ &= (A \otimes \lambda^\top)(\nu_t \otimes X_t) dt + dM_t \\ &= A\nu_t \lambda^\top X_t dt + dM_t. \end{aligned}$$

Using  $\begin{pmatrix} 0 & 1 & \cdots & n \end{pmatrix} A\nu_t = \begin{pmatrix} n & \cdots & 1 & 0 \end{pmatrix} \nu_t = n - N_t$ , we get from the last display the decomposition  $dN_t = (n - N_t)\lambda^\top X_t dt + dm_t$  back.

Letting  $\pi(t) = \mathbb{E}\zeta_t$ , we obtain from Proposition 2.3 the ODE

$$\dot{\pi}(t) = \mathbf{Q}\pi(t) \quad (9)$$

with the initial condition  $\pi(0) = e_0 \otimes x(0)$ , where  $e_0$  has 1 as its first element, all other elements being zero. We will give a rather explicit expression for  $\pi(t) = \exp(\mathbf{Q}t)\pi(0)$ , for which we need some additional results.

The differential equation for  $\pi$  is the following type of forward equation,

$$\dot{F} = \mathbf{Q}F.$$

Here  $F$  can be any matrix valued function of appropriate dimensions. We will block-diagonalize the matrix  $\mathbf{Q}$ . The transformation that is needed for that is given by the matrix  $V$  whose  $ij$ -block  $(i, j = 0, \dots, n)$  is

$$V_{ij} = \binom{n-j}{n-i} (-1)^{i-j} I.$$

Note that  $V_{ij} = 0$  for  $i < j$ ,  $V$  is block lower-triangular. The inverse matrix is also block lower-triangular with blocks

$$V_{ij}^{-1} = \binom{n-j}{n-i} I.$$

One may check by direct computation that indeed  $VV^{-1} = I$ . It is straightforward to verify that  $\mathbf{Q}^V := V^{-1}\mathbf{Q}V$  is block-diagonal with  $i$ -th block ( $i = 0, \dots, n$ ) equal to

$$\mathbf{Q}_i^V = Q_{(n-i)\lambda}.$$

Putting  $G = V^{-1}F$  we obtain

$$\dot{G} = \mathbf{Q}^V G,$$

whose solution satisfying  $G(0) = I$  is block diagonal with  $i$ -th block  $G_i(t) = \exp(Q_{(n-i)\lambda}t)$ . We thus obtain the following lemma.

**Lemma 2.5** *The solution to the forward ODE  $\dot{F} = \mathbf{Q}F$  with initial condition  $F(0)$  is given by  $F(t) = \exp(\mathbf{Q}t)F(0)$ , where*

$$\exp(\mathbf{Q}t) = V \begin{pmatrix} \exp(Q_{n\lambda}t) & & \\ & \ddots & \\ & & \exp(Q_{\lambda}t) \end{pmatrix} V^{-1}.$$

If  $F(t) = \exp(\mathbf{Q}t)$ , its blocks  $F_{ij}(t)$  can be explicitly computed. One has  $F_{ij}(t) = 0$  if  $i < j$ , and for  $i \geq j$  it holds that

$$F_{ij}(t) = \binom{n-j}{n-i} \sum_{k=j}^i (-1)^{i-k} \binom{i-j}{i-k} \exp(Q_{(n-k)\lambda}t).$$

**Proof** We use the block triangular structure of  $V$  and  $V^{-1}$  together with the block diagonal structure of  $\mathbf{Q}^V$  to compute

$$\begin{aligned} F_{ij}(t) &= \sum_{k=j}^i V_{ik} \exp(Q_{(n-k)\lambda}t) V_{kj} \\ &= \sum_{k=j}^i \binom{n-k}{n-i} (-1)^{i-k} \exp(Q_{(n-k)\lambda}t) \binom{n-j}{n-k} \\ &= \binom{n-j}{n-i} \sum_{k=j}^i (-1)^{i-k} \binom{i-j}{i-k} \exp(Q_{(n-k)\lambda}t), \end{aligned}$$

as stated.  $\square$

**Proposition 2.6** *The solution  $\pi(t)$  to the system (9) of ODEs under the initial condition  $\pi(0) = e_0 \otimes x(0)$  has components  $\pi^i(t) \in \mathbb{R}^d$  given by*

$$\pi^i(t) = \binom{n}{i} \sum_{k=0}^i (-1)^{i-k} \binom{i}{i-k} \exp(Q_{(n-k)\lambda}t) x(0). \quad (10)$$

**Proof** We use Lemma 2.5 and recall the specific form of the initial condition  $\pi(0)$ . We have to compute  $\exp(\mathbf{Q}t)\pi(0)$  and obtain from Lemma 2.5 with  $j = 0$  for  $\pi^i(t) = F_{i0}(t)$

$$\begin{aligned} \pi^i(t) &= \binom{n}{n-i} \sum_{k=0}^i (-1)^{i-k} \binom{i}{i-k} \exp(Q_{(n-k)\lambda}t) x(0) \\ &= \binom{n}{i} \sum_{k=0}^i (-1)^{i-k} \binom{i}{k} \exp(Q_{(n-k)\lambda}t) x(0). \end{aligned}$$

□

**Remark 2.7** Let us look at a special case,  $n = 1$ . Then we can write  $N_t = Y_t$  and it is sufficient to compute

$$\pi^1(t) = \mathbb{E}(Y_t X_t) = (\exp(Qt) - \exp(Q_\lambda t))x(0). \quad (11)$$

As a consequence we are able to compute  $\mathbb{P}(Y_t = 1) = \mathbf{1}^\top \mathbb{E}(Y_t X_t)$ ,

$$\mathbb{P}(Y_t = 1) = 1 - \mathbf{1}^\top \exp(Q_\lambda t)x(0),$$

since  $\mathbf{1}^\top \exp(Qt) = \mathbf{1}^\top$ . As  $\exp(Qt) \rightarrow \pi \mathbf{1}^\top$ , we conclude in view of Lemma 2.2 from (11) that  $\pi^1(t) \rightarrow \pi$  for  $t \rightarrow \infty$ . This result should be obvious, as  $Y_t$  eventually becomes 1 and  $X_t$  converges in distribution to its invariant law.

For the case  $n > 1$  the expressions for  $\pi^i(t)$  are a bit complicated, but their asymptotic values for  $t \rightarrow \infty$ , are as expected,  $\pi^i(t) \rightarrow 0$  for  $i < n$ , whereas  $\pi^n(t) \rightarrow \pi$ . This again follows from Lemma 2.2.

Proposition 2.6 has the following corollary.

**Corollary 2.8** *Let  $\phi(t, u) = \mathbb{E} \exp(iu N_t) X_t$ . It holds that*

$$\phi(t, u) = \sum_{k=0}^n \binom{n}{k} \exp(iuk) (1 - \exp(iu))^{n-k} \exp(Q_{(n-k)\lambda} t) x(0).$$

**Proof** We shall use the elementary identity

$$\sum_{k=j}^n \beta^k \binom{n}{k} \binom{k}{j} = \binom{n}{j} \beta^j (1 + \beta)^{n-j}$$

for  $\beta = -e^{-iu}$  in the last step in the chain of equalities below. From Proposition 2.6 we obtain

$$\begin{aligned} \mathbb{E} \exp(iu N_t) X_t &= \sum_{k=0}^n e^{iuk} \pi^k(t) \\ &= \sum_{k=0}^n e^{iuk} \binom{n}{k} \sum_{j=0}^k (-1)^{k-j} \binom{k}{j} \exp(Q_{(n-j)\lambda} t) x(0) \\ &= \sum_{j=0}^n \sum_{k=j}^n (-e^{iu})^k \binom{n}{k} \binom{k}{j} (-1)^j \exp(Q_{(n-j)\lambda} t) x(0) \\ &= \sum_{j=0}^n \binom{n}{j} e^{iju} (1 - e^{iu})^{n-j} \exp(Q_{(n-j)\lambda} t) x(0). \end{aligned}$$

□

**Remark 2.9** Alternatively, one can compute a moment generating function  $\psi(t, v) = \mathbb{E} \exp(-v N_t) X_t$  for  $v \geq 0$ . Let  $B$  have a binomial distribution with parameters  $n$  and  $p = 1 - \exp(-v)$ . Then we have for  $\psi(t, v)$  the compact expression  $\psi(t, v) = \mathbb{E} \exp((Q - B \text{diag}(\lambda))t) x(0) = \mathbb{E} \exp(Q_{\lambda B} t) x(0)$ .

**Remark 2.10** There appears to be no simpler representation for  $\phi(t, u)$ . We note that this function also satisfies the PDE

$$\dot{\phi}(t, u) = (Q + n(e^{iu} - 1)\text{diag}(\lambda))\phi(t, u) + i(e^{iu} - 1)\text{diag}(\lambda)\frac{\partial\phi(t, u)}{\partial u}. \quad (12)$$

Just by computing the partial derivatives, one verifies that this equation holds. Alternatively, one can apply the Itô formula to  $\exp(iuN_t)X_t$  followed by taking expectations.

### 2.2.3 Conditional probabilities

The vehicle we use is the process  $\zeta$ , recall  $\zeta_t = \nu_t \otimes X_t$ . Our aim is to find expressions for  $\zeta_{t|s} = \mathbb{E}[\zeta_t|\mathcal{F}_s]$  for  $t > s$ , from which one can deduce the conditional probabilities  $\mathbb{E}[\nu_t|\mathcal{F}_s]$  and  $\mathbb{E}[N_t|\mathcal{F}_s]$ . By the Markov property, Proposition 2.3, we have  $\mathbb{E}[\zeta_t|\mathcal{F}_s] = \exp(\mathbf{Q}(t-s))\zeta_s$ . Let  $\zeta_{t|s} = \mathbb{E}[\zeta_t|\mathcal{F}_s]$  and  $\zeta_{t|s}^k = \mathbb{E}[\mathbf{1}_{\{N_t=k\}}X_t|\mathcal{F}_s]$ . We aim at a more explicit representation of the conditional probabilities  $\zeta_{t|s}^k$  for  $k \geq 0$ . Note that  $\zeta_{t|s}^k = (e_k^\top \otimes I)\zeta_{t|s}$ . Hence  $\zeta_{t|s}^k = (e_k^\top \otimes I)\exp(\mathbf{Q}(t-s))\zeta_s$ . Using Lemma 2.5, we have

$$\zeta_{t|s}^k = (e_k^\top \otimes I)V \begin{pmatrix} \exp(Q_{n\lambda}(t-s)) & & \\ & \ddots & \\ & & \exp(Q(t-s)) \end{pmatrix} V^{-1}\zeta_s.$$

By matrix computations as before this leads to the following result.

**Proposition 2.11** *It holds that*

$$\zeta_{t|s}^k = \sum_{j=0}^k \binom{n-j}{k-j} \sum_{i=0}^k (-1)^{k-i} \binom{k-j}{k-i} \exp(Q_{(n-i)\lambda}(t-s)) \zeta_s^j.$$

Note that in the formula of this proposition, only one of the  $\zeta_s^j$  is different from zero and then equal to  $X_s$ . Effectively, the sum over  $j$  thus reduces to one term only. The conditional probabilities  $\nu_{t|s}^k = \mathbb{P}(N_t = k|\mathcal{F}_s)$  can now simply be computed as  $\mathbf{1}^\top \zeta_{t|s}^k$ . Note that these still depend on  $X_s$ , and one has the explicit expression

$$\mathbb{E}[\nu_{t|s}^k|\mathcal{F}_s] = \sum_{j=0}^n \binom{n-j}{n-k} \sum_{i=j}^k (-1)^{k-i} \binom{k-j}{k-i} \mathbf{1}^\top \exp(Q_{(n-i)\lambda}(t-s)) X_s \nu_s^j.$$

**Remark 2.12** Consider the special case  $n = 1$  and let  $Z_t = Y_t X_t$ ,  $Y_t$  as in Section 2.1.2. This amounts to taking  $k = n = 1$  in Proposition 2.11 and one gets for  $Z_{t|s} = \mathbb{E}[Z_t|\mathcal{F}_s^Y]$  the simpler expression

$$Z_{t|s} = \exp(Q_\lambda(t-s))Z_s + (\exp(Q(t-s)) - \exp(Q_\lambda(t-s)))X_s. \quad (13)$$

The next purpose is to compute  $\mathbb{E}[e^{iuN_t}X_t|\mathcal{F}_s]$  and from that one  $\mathbb{E}[e^{iuN_t}|\mathcal{F}_s] = \mathbf{1}^\top \mathbb{E}[e^{iuN_t}X_t|\mathcal{F}_s]$ .

**Proposition 2.13** *The following hold.*

$$\begin{aligned}\mathbb{E}[e^{iuN_t} X_t | \mathcal{F}_s] &= \sum_{k=0}^n \sum_{j=k}^n \binom{n-k}{j-k} (1 - e^{iu})^{n-j} e^{iuj} \exp(Q_{(n-j)\lambda}(t-s)) \zeta_s^k, \\ \mathbb{E}[e^{iuN_t} | \mathcal{F}_s] &= \sum_{k=0}^n \sum_{j=k}^n \binom{n-k}{j-k} (1 - e^{iu})^{n-j} e^{iuj} \mathbf{1}^\top \exp(Q_{(n-j)\lambda}(t-s)) \zeta_s^k.\end{aligned}\tag{14}$$

**Proof** We start from the identity  $e^{iuN_t} X_t = \mathbf{F} \zeta_t$ , with  $\mathbf{F} = e(u) \otimes I$ , where  $e(u) = \begin{pmatrix} 1 & e^{iu} & \dots & e^{niu} \end{pmatrix}$ . Hence we have

$$\mathbb{E}[e^{iuN_t} X_t | \mathcal{F}_s] = (e(u) \otimes I) \exp(\mathbf{Q}(t-s)) \zeta_s.$$

This can be put into the asserted more explicit representation, involving the matrices  $Q_{k\lambda}$  by application of Proposition 2.11. The second assertion is a trivial consequence.  $\square$

It is conceivable that only  $N$  is observed, and not the background process  $X$ . In such a case one is only able to compute conditional expectation of quantities as above conditioned on  $\mathcal{F}_s^N$  instead of  $\mathcal{F}_s$ . See Section 4.1 for results.

### 3 The Markov Modulated Poisson process

In this section we study MM Poisson processes. These have an intensity process  $\lambda_t = \lambda^\top X_t$ , using the same notation as before. In terms of defaultable obligors, such processes occur as limits of the total number of defaults  $N_t$  as in Section 2.2 where  $n \rightarrow \infty$  and the vector  $\lambda$  is scaled to become  $\lambda/n$ , as we shall see later. So we can use this to approximate the total number of defaults in a market with a large number of obligors, where each of them has small default rate.

#### 3.1 The model

The point of departure is to postulate the dynamics of the counting process  $N$  as

$$dN_t = \lambda^\top X_t dt + dm_t.$$

We follow the same approach as before. So we use that conditionally on  $\mathcal{F}^X$  we have that  $N_t$  has a Poisson( $\Lambda_t$ ) distribution with  $\Lambda_t = \int_0^t \lambda^\top X_s ds$ . It follows that

$$\mathbb{E}[\mathbf{1}_{\{N_t=k\}} X_t | \mathcal{F}^X] = \frac{1}{k!} \Lambda_t^k \exp(-\Lambda_t) X_t =: p^k(t) X_t,$$

and

$$\frac{d}{dt} p^k(t) = p^{k-1}(t) - p^k(t) \lambda^\top X_t.$$

Then we obtain

$$d\mathbb{E}[\mathbf{1}_{\{N_t=k\}} X_t | \mathcal{F}^X] = (p^{k-1}(t) - p^k(t)) \text{diag}(\lambda) X_t dt + p^k(t) (Q X_t dt + dM_t),$$

and with  $\pi^k(t) = \mathbb{E}(p^k(t) X_t)$  we find

$$\dot{\pi}^k(t) = \text{diag}(\lambda) \pi^{k-1}(t) + (Q - \text{diag}(\lambda)) \pi^k(t).$$

For  $k = 0$ , one immediately finds the solution  $\pi^0(t) = \exp(Q_\lambda t)x(0)$ . For  $k > 0$  there seems to be no simply expression in terms of exponential of  $Q$  and  $Q_{k\lambda}$  as in Proposition 2.6, not even for  $k = 1$ , although one has

$$\pi^1(t) = \int_0^t \exp(-Q_\lambda(t-s)) \text{diag}(\lambda) \exp(Q_\lambda s) \, ds \, x(0).$$

However, it is possible to get a formula for the vector

$$\Pi^n(t) = \begin{pmatrix} \pi^0(t) \\ \vdots \\ \pi^n(t) \end{pmatrix},$$

since it satisfies the ODE

$$\dot{\Pi}^n(t) = \mathbf{Q}_n \Pi^n(t),$$

where  $\mathbf{Q}_n \in \mathbb{R}^{(n+1)d \times (n+1)d}$  is given by

$$\mathbf{Q}_n = \begin{pmatrix} Q - \text{diag}(\lambda) & 0 & \cdots & \cdots & 0 \\ \text{diag}(\lambda) & Q - \text{diag}(\lambda) & 0 & & 0 \\ 0 & \text{diag}(\lambda) & \ddots & \ddots & \vdots \\ \vdots & & \ddots & Q - \text{diag}(\lambda) & 0 \\ 0 & \cdots & 0 & \text{diag}(\lambda) & Q - \text{diag}(\lambda) \end{pmatrix}.$$

Together with the initial conditions  $\pi^k(0) = \delta_{k0}x(0)$ , one obtains

$$\Pi^n(t) = \exp(\mathbf{Q}_n t)(e_0^n \otimes x(0)),$$

where  $e_0^n$  is the first basis vector of  $\mathbb{R}^{n+1}$ . An elementary expression for  $\exp(\mathbf{Q}_n t)$  is not available due to the fact that  $Q - \text{diag}(\lambda)$  and  $\text{diag}(\lambda)$  do not commute. Besides,  $\mathbf{Q}_n$  is block lower triangular with identical blocks on the main diagonal and therefore cannot be block diagonalized.

However, in the present case there is a nice expression for the characteristic function  $\phi(t, u) = \mathbb{E} \exp(iu N_t) X_t$ , unlike the situation of Corollary 2.8. To determine  $\phi(t, u)$ , we apply the Itô formula (note that  $[N, X] = 0$ ) and obtain

$$d \exp(iu N_t) X_t = (e^{iu} - 1) e^{iu N_t -} X_{t-} dN_t + e^{iu N_t -} dX_t, \quad (15)$$

which yields after taking expectations and using the dynamics of  $X$  and  $N$

$$\dot{\phi}(t, u) = ((e^{iu} - 1) \text{diag}(\lambda) + Q) \phi(t, u).$$

Hence

$$\phi(t, u) = \exp(((e^{iu} - 1) \text{diag}(\lambda) + Q)t) x(0).$$

Contrary to the  $\pi^k(t)$  of Proposition 2.6 we thus found a *simple* formula for  $\phi(t, u)$ . This formula is in line with [2, Proposition 1.6] for Markovian arrival processes.

**Remark 3.1** It is possible to obtain the above results as limits from results in Section 2.2.2, by replacing there  $\lambda$  by  $\lambda/n$  and letting  $n \rightarrow \infty$ .

If we look at the moment generating functions  $\psi(t, v) = \mathbb{E} \exp(-vN_t)X_t$ , we have  $\psi(t, v) = \exp((Q - (1 - e^{-v})\text{diag}(\lambda))t)x(0)$ . Replace in Remark 2.9 the parameter  $\lambda$  with  $\lambda/n$  and let  $n \rightarrow \infty$  and write  $B_n$  instead of  $B$ . Then we have  $\psi_n(t, v) = \mathbb{E} \exp((Q - \text{diag}(\lambda)B_n/n)t)x(0)$ . As  $B_n/n \rightarrow 1 - e^{-v}$  a.s., we obtain  $\exp((Q - \text{diag}(\lambda)B_n/n)t) \rightarrow \exp((Q - \text{diag}(\lambda)(1 - e^{-v}))t)$  a.s. Since the exponentials are bounded, we also have convergence of the expectations by dominated convergence. Replacing  $-v$  with  $iu$  gives the characteristic function.

### 3.2 Conditional probabilities

Mimicking the approach of Section 2.2.2, we consider again the  $\nu_t^k = \mathbf{1}_{\{N_t=k\}}$ . Let

$$\bar{\nu}_t^n = \begin{pmatrix} \nu_t^0 \\ \vdots \\ \nu_t^n \end{pmatrix}.$$

Then  $\bar{\nu}^n$  still satisfies Equation (6). Combining this with the dynamics of  $N$ , we obtain the semimartingale decomposition

$$d\bar{\nu}_t^n = \lambda^\top X_t (J - I) \bar{\nu}_t^n dt + dM_t.$$

Letting  $\bar{\zeta}_t^n = \bar{\nu}_t^n \otimes X_t$ , then we can derive, similar to the approach of Section 2.2.2,

$$d\bar{\zeta}_t^n = \mathbf{Q}_n \bar{\zeta}_t^n dt + dM_t.$$

This is for each  $n$  a finite dimensional system, which can be extended to an infinite dimensional system for  $\zeta_t$ . The resulting infinite coefficient matrix will be lower triangular again,

$$d\zeta_t = \mathbf{Q}_\infty \zeta_t dt + dM_t,$$

where  $\mathbf{Q}_\infty = I_\infty \otimes Q_\lambda - J_\infty \otimes \text{diag}(\lambda)$  with  $I_\infty$  the infinite dimensional identity matrix and  $J_\infty$  the infinite dimensional counterpart of the earlier encountered matrix  $J$ . It follows that for the vector of conditional probabilities we have

$$\mathbb{E}[\zeta_t | \mathcal{F}_s] = \exp(\mathbf{Q}_\infty(t - s)) \bar{\zeta}_s.$$

This looks like an infinite dimensional expression, but  $\mathbb{E}[\mathbf{1}_{\{N_t=n\}} X_t | \mathcal{F}_s]$  can be computed from  $\mathbb{E}[\bar{\zeta}_t^n | \mathcal{F}_s] = \exp(\mathbf{Q}_n(t - s)) \bar{\zeta}_s^n$ , which effectively reduces the infinite dimensional system to a finite dimensional one. One can now also compute, with  $\ell_n^\top = (0 \ \cdots \ 0 \ 1) \in \mathbb{R}^{1 \times (n+1)}$ ,

$$\mathbb{P}(N_t = n, X_t = e_j | \mathcal{F}_s) = (\ell_n^\top \otimes e_j^\top) \exp(\mathbf{Q}_n(t - s)) \bar{\zeta}_s^n.$$

### 3.3 Conditional characteristic function

Our aim is to find an expression for  $\phi_{t|s} := \mathbb{E}[\exp(iuN_t)X_t | \mathcal{F}_s]$ . Since we deal in the present section with the MM Poisson process  $N$ , the bivariate process  $(X, N)$ , unlike its counterpart in Section 2, is an instance of a Markov additive process [2], and  $\mathbb{E}[\exp(iu(N_t - N_s))X_t | \mathcal{F}_s]$  will only depend on  $X_s$ . We first follow the forward approach.

**Proposition 3.2** *It holds that*

$$\phi_{t|s} = \exp \left( ((e^{iu} - 1)\text{diag}(\lambda) + Q)(t - s) \right) e^{iuN_s} X_s. \quad (16)$$

**Proof** Starting point is Equation (15). We use the dynamics of  $N$  and  $X$  to get the semimartingale decomposition

$$\begin{aligned} d \exp(iuN_t) X_t &= (e^{iu} - 1) e^{iuN_t} \text{diag}(\lambda) X_t dt + e^{iuN_t} Q X_t dt + dM_t \\ &= ((e^{iu} - 1)\text{diag}(\lambda) + Q) e^{iuN_t} X_t dt + dM_t. \end{aligned}$$

Let  $t \geq s$ . We obtain (differentials w.r.t.  $t$ )

$$d\phi_{t|s} = ((e^{iu} - 1)\text{diag}(\lambda) + Q)\phi_{t|s} dt,$$

which has the desired solution.  $\square$

Next we outline the backward approach. Observe first that  $\phi_{t|s}$  is a martingale in the  $s$ -parameter and that due to the fact that  $(N, X)$  is Markov, we can write for some function  $\Phi$ ,  $\phi_{t|s} = \Phi(t - s, N_s) X_s$ . We identify  $\Phi$  as follows, using the Itô formula w.r.t.  $s$ . We obtain

$$\begin{aligned} d\phi_{t|s} &= \left( -\dot{\Phi}(t - s, N_s) ds + (\Phi(t - s, N_{s-} + 1) - \Phi(t - s, N_{s-})) dN_s \right) X_{s-} \\ &\quad + \Phi(t - s, N_{s-}) dX_s \\ &= \left( -\dot{\Phi}(t - s, N_s) + (\Phi(t - s, N_s + 1) - \Phi(t - s, N_s)) \text{diag}(\lambda) \right) X_s ds \\ &\quad + \Phi(t - s, N_s) Q X_s ds + dM_s. \end{aligned}$$

The above mentioned martingale property leads to the system of ODEs ( $n \geq 0$ )

$$\dot{\Phi}(t, n) = \Phi(t, n + 1) \text{diag}(\lambda) + \Phi(t, n) (Q - \text{diag}(\lambda)). \quad (17)$$

We have the initial conditions  $\Phi(0, n) = \exp(iun)$ . To know  $\Phi(t, n)$  it seems necessary to know  $\Phi(t, n + 1)$ , which suggest that the ODEs are difficult to solve constructively. Instead, we pose a solution, we will verify that

$$\Phi(t, n) = \exp \left( ((e^{iu} - 1)\text{diag}(\lambda) + Q)t \right) e^{iun}.$$

Differentiation of the given expression for  $\Phi(t, n)$  gives

$$\dot{\Phi}(t, n) = \Phi(t, n) ((e^{iu} - 1)\text{diag}(\lambda) + Q).$$

Note that  $\Phi(t, n + 1) = \Phi(t, n) e^{iu}$ . Insertion of this into the ODE gives

$$\dot{\Phi}(t, n) = \Phi(t, n) (e^{iu} \text{diag}(\lambda) + (Q - \text{diag}(\lambda))),$$

which coincides with (17).

## 4 Filtering

Let  $N$  be a counting process with predictable intensity process  $\lambda$ . In many cases it is conceivable that  $\lambda$  is an unobserved process and expressions in terms of  $\lambda$  are



not always useful. Let  $\hat{\lambda}_t = \mathbb{E}[\lambda_t | \mathcal{F}_t^N]$ . Then the semimartingale decomposition of  $N$  w.r.t. the filtration  $\mathbb{F}^N$  is given by

$$dN_t = \hat{\lambda}_t dt + d\hat{m}_t,$$

where  $\hat{m}$  is a (local) martingale w.r.t.  $\mathbb{F}^N$ . The general filter of the Markov chain  $X$ ,  $\hat{X}_t = \mathbb{E}[X_t | \mathcal{F}_t^N]$  satisfies the following well known formula (see [7], originating from [28]) with  $Q$  as in Section 2.1.2

$$d\hat{X}_t = Q\hat{X}_t dt + \hat{\lambda}_{t-}^+ (\widehat{X\lambda}_{t-} - \hat{X}_{t-}\hat{\lambda}_{t-})(dN_t - \hat{\lambda}_t dt),$$

where  $\widehat{X\lambda}_t = \mathbb{E}[X_t \lambda_t | \mathcal{F}_t^N]$  and where we use the notation  $x^+ = \mathbf{1}_{x \neq 0}/x$  for a real number  $x$ . For any of the previously met models for the counting process  $N$  we have a predictable intensity process of the form  $\lambda_t = \lambda^\top X_{t-} f(N_{t-})$ , where  $f$  depends on the specific model at hand. It follows that  $\hat{\lambda}_t = \lambda^\top \hat{X}_{t-} f(N_{t-})$ . In all cases we consider it happens that  $f(N_t)$  remains zero after it has reached zero, and hence  $N$  stops jumping as soon as  $f(N_t) = 0$ . Since  $\lambda^\top X_t > 0$ , with the convention  $\frac{0}{0} = 0$  the above filter equation reduces to

$$d\hat{X}_t = Q\hat{X}_t dt + \frac{1}{\lambda^\top \hat{X}_{t-}} (\text{diag}(\lambda)\hat{X}_{t-} - \hat{X}_{t-}\lambda^\top \hat{X}_{t-})(dN_t - \hat{\lambda}_t dt). \quad (18)$$

For the specific models we have encountered we give in the next sections more results on  $\hat{X}$ .

#### 4.1 Filtering for the MM multiple point process

The notation of this section is as in Section 2.2.2 and subsequent sections. Let  $\hat{\zeta}_t = \mathbb{E}[\zeta_t | \mathcal{F}_t^N]$ . Then  $\hat{\zeta}_t = \nu_t \otimes \hat{X}_t$ , where  $\hat{X}_t = \mathbb{E}[X_t | \mathcal{F}_t^N]$ . For  $\hat{X}_t$  we have from (18),

$$d\hat{X}_t = Q\hat{X}_t dt + \frac{1}{\lambda^\top \hat{X}_{t-}} \left( \text{diag}(\lambda)\hat{X}_{t-} - \hat{X}_{t-}\hat{X}_{t-}^\top \lambda \right) (dN_t - (n - N_t)\lambda^\top \hat{X}_t dt).$$

At the jump times  $\tau_k$  ( $k = 1, \dots, n$ ) (these are the order statistics of the original default times  $\tau^i$ ) of  $N$  we thus have

$$X_{\tau_k} = \frac{1}{\lambda^\top \hat{X}_{\tau_k-}} \text{diag}(\lambda)\hat{X}_{\tau_k-}$$

Between the jump times,  $\hat{X}$  evolves according to the ODE

$$\frac{d\hat{X}_t}{dt} = Q\hat{X}_t - (n - N_t)(\text{diag}(\lambda)\hat{X}_{t-} - \hat{X}_{t-}\hat{X}_{t-}^\top \lambda),$$

which is also valid after the last jump of  $N$ . It follows that for  $t \geq \tau_n$  we have  $\hat{X}_t = \exp(Q(t - \tau_n))\hat{X}_{\tau_n}$ .

Below we need  $[\nu, \hat{X}]_t^\otimes = \sum_{s \leq t} \Delta \nu_s \otimes \Delta \hat{X}_s$ . Using the equations for  $\nu$  and  $\hat{X}$ , we find

$$d[\nu, \hat{X}]_t^\otimes = \frac{1}{\lambda^\top \hat{X}_{t-}} ((J - I) \otimes (\text{diag}(\lambda) - \lambda^\top \hat{X}_{t-} I)) \hat{\zeta}_{t-} dN_t.$$

For  $\hat{\zeta}_t$  we have, using the product formula for tensors,

$$d\hat{\zeta}_t = d\nu_t \otimes \hat{X}_{t-} + \nu_{t-} \otimes d\hat{X}_t + d[\nu, \hat{X}]_t^{\otimes}.$$

This yields after some tedious computations the following semimartingale decomposition for  $\hat{\zeta}$

$$\begin{aligned} d\hat{\zeta}_t &= (I \otimes Q + (n - N_t)(J - I) \otimes \text{diag}(\lambda)) \hat{\zeta}_t dt \\ &\quad + \frac{1}{\lambda^\top \hat{X}_{t-}} (J \otimes \text{diag}(\lambda) - \lambda^\top \hat{X}_{t-} I \otimes I) \hat{\zeta}_{t-} d\hat{m}_t \\ &= \mathbf{Q} \hat{\zeta}_t dt + \frac{1}{\lambda^\top \hat{X}_{t-}} (J \otimes \text{diag}(\lambda) - \lambda^\top \hat{X}_{t-} I \otimes I) \hat{\zeta}_{t-} d\hat{m}_t, \end{aligned}$$

where  $d\hat{m}_t = dN_t - (n - N_t)\lambda^\top \hat{X}_t dt$  and  $\mathbf{Q}$  as in Section 2.2.2.

Here are two applications. One can now compute

$$\mathbb{P}(N_t = k | \mathcal{F}_s^N) = \mathbf{1}^\top \mathbb{E}[\zeta_{t|s}^k | \mathcal{F}_s^N] = \mathbf{1}^\top \hat{\zeta}_{t|s}^k,$$

for which we can use  $\hat{\zeta}_{t|s} = \exp(\mathbf{Q}(t - s))\hat{\zeta}_s$ . Formula (14) yields for the conditional characteristic function of  $N_t$  given its own past until time  $s < t$  the explicit expression

$$\mathbb{E}[e^{iuN_t} | \mathcal{F}_s^N] = \sum_{k=0}^n \sum_{j=k}^n \binom{n-k}{j-k} (1 - e^{iu})^{n-j} e^{iuj} \mathbf{1}^\top \exp(Q_{(n-j)\lambda}(t - s)) \hat{X}_s \nu_s^k.$$

In case  $n = 1$  the above formulas simplify considerably. Here are a few examples, where we use the notation of Section 2.1.2. Suppose that only  $Y$  is observed. Let  $\mathcal{F}_t^Y = \sigma(Y_s, 0 \leq s \leq t)$ . With  $Z_t := Y_t X_t$  we want to compute  $\hat{Z}_{t|s} := \mathbb{E}[Z_t | \mathcal{F}_s^Y]$  for  $t \geq s$ . Let  $\hat{X}_t = \mathbb{E}[X_t | \mathcal{F}_t^Y]$ , then obviously,  $\hat{Z}_{t|s} = \hat{X}_{t|s} Y_s$ . Moreover, one has from (13)

$$\hat{Z}_{t|s} = \exp(Q(t - s))\hat{X}_s - \exp(Q_\lambda(t - s))\hat{X}_s(1 - Y_s).$$

As a consequence we have for  $\hat{Y}_{t|s} = \mathbf{1}^\top \hat{Z}_{t|s}$

$$\hat{Y}_{t|s} = 1 - \mathbf{1}^\top \exp(Q_\lambda(t - s))\hat{X}_s(1 - Y_s).$$

## 4.2 Filtering for the MM Poisson process

The filter equations now take the familiar form

$$d\hat{X}_t = Q\hat{X}_t dt + \frac{1}{\lambda^\top \hat{X}_{t-}} \left( \text{diag}(\lambda)\hat{X}_{t-} - \hat{X}_{t-}\hat{X}_{t-}^\top \lambda \right) (dN_t - \lambda^\top \hat{X}_t dt).$$

For  $\bar{\nu}_t$  we have the infinite dimensional analogue of (6). This leads for  $\hat{\zeta}_t = \bar{\nu}_t \otimes \hat{X}_t$  as in a Section 4.1 to

$$d\hat{\zeta}_t = \mathbf{Q}_\infty \hat{\zeta}_t dt + \frac{1}{\lambda^\top \hat{X}_{t-}} (J_\infty \otimes \text{diag}(\lambda) - \lambda^\top \hat{X}_{t-} I_\infty \otimes I_\infty) \hat{\zeta}_{t-} (dN_t - \lambda^\top \hat{X}_t dt).$$

Note that this system is infinite dimensional, but for each  $n$  we also have for  $\hat{\zeta}_t^n = \mathbb{E}[\hat{\zeta}_t^n | \mathcal{F}_t^N]$  the truncated finite dimensional system

$$d\hat{\zeta}_t^n = \mathbf{Q}_n \hat{\zeta}_t^n dt + \frac{1}{\lambda^\top \hat{X}_{t-}} (J \otimes \text{diag}(\lambda) - \lambda^\top \hat{X}_{t-} I \otimes I) \hat{\zeta}_{t-}^n (dN_t - \lambda^\top \hat{X}_t dt).$$

For the conditional characteristic function  $\mathbb{E}[\exp(iuN_t)X_t | \mathcal{F}_s^N]$  we have

$$\mathbb{E}[\exp(iuN_t)X_t | \mathcal{F}_s^N] = \exp(((e^{iu} - 1)\text{diag}(\lambda) + Q)(t - s)) e^{iuN_s} \hat{X}_s,$$

whereas  $\psi_t = e^{iuN_t} \hat{X}_t$  satisfies the equation  $(d\hat{m}_t = dN_t - \lambda^\top \hat{X}_t dt)$

$$d\psi_t = \left( \frac{e^{iu}}{\lambda^\top \hat{X}_{t-}} \text{diag}(\lambda) - I \right) \psi_{t-} d\hat{m}_t + (Q + (e^{iu} - 1)\text{diag}(\lambda)) \psi_t dt.$$

## 5 Rapid switching

In this section we present some auxiliary results that we shall use in obtaining limits for the various default processes when the Markov chain evolves under a rapid switching regime, i.e. the transition matrix  $Q$  will be replaced with  $\alpha Q$ , where  $\alpha > 0$  tends to infinity. In the first two results and their proofs we use the notation  $C(M)$  for the matrix of cofactors of a square matrix  $M$ . Throughout this section we write  $\lambda_\infty$  for  $\lambda^\top \pi$ .

**Lemma 5.1** *Let  $Q$  have a unique invariant vector  $\pi$ . Then*

$$C(Q) = q \pi \mathbf{1}^\top,$$

where the constant  $q$  can be computed as  $\det(\hat{Q})$ , where  $\hat{Q}$  is obtained from  $Q$  by replacing its last row with  $\mathbf{1}^\top$ .

**Proof** Note first that  $\pi$  can be obtained as the solution to  $\hat{Q}\pi = e_d$ , where  $e_d$  is the last basis vector of  $\mathbb{R}^d$ . By Cramer's rule  $\pi$  can be expressed using the cofactors of  $\hat{Q}$ . In particular,  $\pi_d = \hat{C}_{dd} / \det(\hat{Q})$ , where  $\hat{C}$  is the cofactor matrix of  $\hat{Q}$ . But  $\hat{C}_{dd} = C_{dd}$ , so  $\pi_d = C_{dd} / \det(\hat{Q})$ .

Write  $C = C(Q)$  and recall that  $CQ = \det(Q)$  and hence zero. It follows that every row of  $C$  is a left eigenvector of  $Q$ . Since  $Q$  has rank  $d - 1$  by its assumed irreducibility, every row of  $C$  is a multiple of  $\mathbf{1}^\top$ . Hence  $C = \alpha \mathbf{1}^\top$ , for some  $\alpha \in \mathbb{R}^{d \times 1}$ . By similar reasoning,  $C = \pi \beta$  for some  $\beta \in \mathbb{R}^{1 \times d}$ . We conclude that  $C = q \pi \mathbf{1}^\top$  for some real constant  $q$ . Use now  $C_{dd} = q \pi_d$  and the above expression for  $\pi_d$  to arrive at  $q = \det(\hat{Q})$ .  $\square$

**Proposition 5.2** *Let  $Q$  have a unique invariant vector  $\pi$  and let all  $\lambda_i$  be positive. Then  $(\alpha Q - \text{diag}(\lambda))^{-1} \rightarrow -\frac{\pi \mathbf{1}^\top}{\lambda_\infty}$  for  $\alpha \rightarrow \infty$ .*

**Proof** We have seen in Section 2.1.2 that  $Q - \text{diag}(\lambda)$  is invertible if all  $\lambda_i > 0$  and so the same is true for  $\alpha Q - \text{diag}(\lambda)$ . Both  $\det(\alpha Q - \text{diag}(\lambda))$  and the cofactor matrix of  $\alpha Q - \text{diag}(\lambda)$  are polynomials in  $\alpha$  and we compute the leading term. The determinant is computed by summing products of elements of  $\alpha Q - \text{diag}(\lambda)$ , from each row and each column one. The  $\alpha^d$  term in this determinant has coefficient  $\det(Q)$ , which is zero. Consider the term with  $\alpha^{d-1}$ . It is seen to be

equal to  $-\sum_{i=1}^d \lambda_i C(\alpha Q - \text{diag}(\lambda))_{ii} = -\alpha^{d-1} \sum_{i=1}^d \lambda_i C(Q - \text{diag}(\lambda/\alpha))_{ii}$ . For the cofactor matrix itself a similar procedure applies. We get  $C(\alpha Q - \text{diag}(\lambda)) = \alpha^{d-1} C(Q - \text{diag}(\lambda/\alpha))$  and it results from Lemma 5.1 that for  $\alpha \rightarrow \infty$

$$\frac{C(\alpha Q - \text{diag}(\lambda))}{\det(\alpha Q - \text{diag}(\lambda))} \rightarrow \frac{C(Q)}{-\sum_{i=1}^d \lambda_i C(Q)_{ii}} = -\frac{q\pi \mathbf{1}^\top}{q \sum_{i=1}^n \lambda_i \pi_i} = -\frac{\pi \mathbf{1}^\top}{\lambda_\infty}.$$

□

**Proposition 5.3** *For  $\alpha \rightarrow \infty$  it holds that*

$$\exp((\alpha Q - \text{diag}(\lambda))t) \rightarrow \exp(-\lambda_\infty t) \pi \mathbf{1}^\top.$$

**Proof** For any analytic function  $f : \mathbb{C} \rightarrow \mathbb{C}$ ,  $f(z) = \sum_{k=0}^{\infty} a_k z^k$ , one defines  $f(M) := \sum_{k=0}^{\infty} a_k M^k$  for  $M \in \mathbb{C}^{d \times d}$  (assuming that the power series converges on the spectrum of  $M$ ). It then holds (see also Higham [17, Definition 1.11], where this is taken as a definition of  $f(M)$ ) that

$$f(M) = \frac{1}{2\pi i} \oint_{\Gamma} (zI - M)^{-1} f(z) dz,$$

where  $\Gamma$  is a closed contour such that all eigenvalues of  $M$  are inside it. Take  $M = \alpha Q - \text{diag}(\lambda)$ . It follows from Proposition 5.2, note that also  $\lambda_\infty$  lies inside  $\Gamma$  as it is a convex combination of the  $\lambda_i$ , that  $(zI - \alpha Q + \text{diag}(\lambda))^{-1} \rightarrow \frac{1}{z + \lambda_\infty} \pi \mathbf{1}^\top$ . Hence

$$f(\alpha Q - \text{diag}(\lambda)) \rightarrow \pi \mathbf{1}^\top f(-\lambda_\infty).$$

Apply this to  $f(z) = \exp(tz)$ . □

## 5.1 Rapid switching for the MM multiple point process

Suppose we scale the  $Q$  matrix with  $\alpha \geq 0$ , and we let  $X^\alpha$  be Markov with transition matrix  $\alpha Q$ . Many (random) variables below will be indexed by  $\alpha$  as well. Here is a way to get accelerated dynamics for  $N_t^\alpha$  (previously denoted  $N_t$ ).

Suppose that one takes the original Markov chain  $X$  and replaces the dynamics of  $N$  with one in which  $X$  is accelerated,

$$N_t^\alpha = \int_0^t (n - N_s^\alpha) \lambda^\top X_{\alpha s} ds + m_t. \quad (19)$$

Indeed the process  $X^\alpha$  defined by  $X_t^\alpha = X_{\alpha t}$  has intensity matrix  $\alpha Q$ , and its invariant measure is  $\pi$  again. Recall that, conditionally on  $\mathcal{F}^X$ ,  $N_t^\alpha$  has a  $\text{Bin}(n, 1 - \exp(-\int_0^t \lambda^\top X_{\alpha s} ds))$  distribution and that its unconditional distribution is  $\text{Bin}(n, 1 - \mathbb{E} \exp(-\int_0^t \lambda^\top X_{\alpha s} ds))$ .

The ergodic property of  $X$  gives  $\int_0^t X_{\alpha s} ds = \frac{1}{\alpha} \int_0^{\alpha t} X_s ds \rightarrow \pi t$  a.s. and hence by dominated convergence for the expectations we have that the limit distribution of  $N_t^\alpha$  for  $\alpha \rightarrow \infty$  is  $\text{Bin}(n, 1 - \exp(-\lambda_\infty t))$ . One immediately sees that the default times  $\tau^{\alpha, k}$  convergence in distribution to  $\tau^k$  that are independent and have an exponential distribution with parameter  $\lambda_\infty$ . Keeping this in mind, the other results in this section are easily understandable.

We recall the content of Proposition 5.3. Replacing  $\lambda$  with  $k\lambda$  for  $k \geq 0$  (zero included) yields

$$\exp((\alpha Q - k \text{diag}(\lambda))t) \rightarrow \exp(-k\lambda_\infty t) \pi \mathbf{1}^\top. \quad (20)$$

To express the dependence of the matrix  $\mathbf{Q}$  given by (8) on  $\alpha$  in the present section, we write  $\mathbf{Q}^\alpha$  (so  $\mathbf{Q}^\alpha = A \otimes \text{diag}(\lambda) + I \otimes \alpha Q$ ) and  $F^\alpha(t)$  instead of  $F(t)$  as given in Lemma 2.5.

**Lemma 5.4** *The solution  $F^\alpha$  to the equation  $\dot{F} = \mathbf{Q}^\alpha F$ , has for  $\alpha \rightarrow \infty$  limit  $F^\infty$  given by its blocks*

$$F_{ij}^\infty(t) = f_{ij}^\infty(t) \pi \mathbf{1}^\top,$$

where the  $f_{ij}^\infty(t)$  are the binomial probabilities on  $n-i$  ‘successes’ of a  $\text{Bin}(n-j, \exp(-\lambda_\infty t))$  distribution,

$$f_{ij}^\infty(t) = \binom{n-j}{n-i} \exp(-(n-i)\lambda_\infty t) (1 - \exp(-\lambda_\infty t))^{i-j}.$$

**Proof** We depart from Lemma 2.5 and the expression for  $F_{ij}^\alpha(t)$  given there when we replace  $Q$  with  $\alpha Q$ . Taking limits for  $\alpha \rightarrow \infty$  yields

$$\begin{aligned} F_{ij}^\infty(t) &= \binom{n-j}{n-i} \sum_{k=j}^i (-1)^{i-k} \binom{i-j}{i-k} \exp(-(n-k)\lambda_\infty t) \pi \mathbf{1}^\top \\ &= \binom{n-j}{n-i} (-1)^{i-j} \exp(-(n-j)\lambda_\infty t) \sum_{l=0}^{i-j} \binom{i-j}{l} (-\exp(\lambda_\infty t))^l \pi \mathbf{1}^\top \\ &= \binom{n-j}{n-i} \exp(-(n-i)\lambda_\infty t) (1 - \exp(-\lambda_\infty t))^{i-j} \pi \mathbf{1}^\top, \end{aligned}$$

from which the assertion follows.  $\square$

**Remark 5.5** One can also use this proposition to show that  $N_t^\alpha$  in the limit has the  $\text{Bin}(n, 1 - \exp(-\lambda_\infty t))$  distribution. Indeed, since  $\nu_0^i = \delta_{i0}$ , we get  $\mathbb{P}(N_t^\alpha = i, X_t = e_j) \rightarrow F_{i0}^\infty(t) = f_{i0}^\infty(t) \pi$  and hence  $\mathbb{P}(N_t^\alpha = i) \rightarrow f_{i0}^\infty(t)$ .

For conditional probabilities one has the following result.

**Corollary 5.6** *Let  $N$  be a process like in Equation (4), with  $\lambda$  replaced with  $\lambda_\infty$ . For  $\alpha \rightarrow \infty$  one has in the limit  $\zeta_{t|s}^i = 0$  for  $i < N_s$  and for  $i \geq N_s$*

$$\zeta_{t|s}^i = \binom{n-N_s}{n-i} \exp(-(n-i)\lambda_\infty(t-s)) (1 - \exp(-\lambda_\infty(t-s)))^{i-N_s} \pi.$$

*It follows that, conditional on  $\mathcal{F}_s$ ,  $N_t - N_s$  has a  $\text{Bin}(n - N_s, 1 - \exp(-\lambda_\infty(t-s)))$  distribution. In fact, one has weak convergence of the  $N^\alpha$  to  $N$ .*

**Proof** We compute in the limit  $\zeta_{t|s}^i = \mathbb{E}[\nu_t^i X_t | \mathcal{F}_s]$  and obtain from Lemma 5.4

$$\begin{aligned}
\zeta_{t|s}^i &= \sum_{j=0}^n F_{ij}^\infty(t-s) \zeta_s^j \\
&= \sum_{j=0}^n f_{ij}^\infty(t-s) \nu_s^j \pi \\
&= \sum_{j=0}^i \binom{n-j}{n-i} \exp(-(n-i)\lambda_\infty(t-s))(1 - \exp(-\lambda_\infty(t-s)))^{i-j} \nu_s^j \pi \\
&= \binom{n-N_s}{n-i} \exp(-(n-i)\lambda_\infty(t-s))(1 - \exp(-\lambda_\infty(t-s)))^{i-N_s} \pi,
\end{aligned}$$

from which the first assertion follows.

Weak convergence can be proved in many ways. Let us first look at the case of one obligor,  $n = 1$ . The integral in Equation (19) is, with  $\tau^\alpha = \tau^{1,\alpha}$  equal to

$$\frac{1}{\alpha} \int_0^{\alpha(\tau^\alpha \wedge t)} \lambda^\top X_u du.$$

Replacing the upper limit of the integral by  $t$ , this almost surely converges to  $\lambda_\infty t$  for  $\alpha \rightarrow \infty$ . In fact this convergence is a.s. uniform. Having already established the convergence in distribution of the  $\tau^\alpha$ , and by switching to an auxiliary space on which the  $\tau^\alpha$  a.s. converge to  $\tau^\infty$ , we get

$$\frac{1}{\alpha} \int_0^{\alpha(\tau^\alpha \wedge t)} \lambda^\top X_u du \rightarrow \int_0^{\alpha(\tau^\infty \wedge t)} \lambda^\top X_u du.$$

This is sufficient, see [24] or [20, Section VIII.3d] to conclude the weak convergence result for the case  $n = 1$ .

For the general case, one first notices that the process  $N^\alpha$  is a sum of MM one point processes that are conditionally independent given  $\mathcal{F}^X$  and become independent in the limit. Combine this with the result for  $n = 1$ . Alternatively, one could apply the results in [20, Section VII.3d] again, although the computations will now be more involved.  $\square$

## 5.2 Rapid switching for the MM Poisson process

As before we replace  $Q$  with  $\alpha Q$  and let  $\alpha \rightarrow \infty$  and denote  $N^\alpha$  the corresponding counting process. We apply Proposition 5.3 to the matrix exponential  $\exp(((e^{iu} - 1)\text{diag}(\lambda) + \alpha Q)(t-s))$ , and we find that the limit for  $\alpha \rightarrow \infty$  equals  $\exp((e^{iu} - 1)\lambda_\infty(t-s))\pi \mathbf{1}^\top$ . Hence, by virtue of (16), we obtain  $\mathbb{E}[\exp(iu N_t^\alpha) X_t | \mathcal{F}_s] \rightarrow \exp((e^{iu} - 1)\lambda_\infty(t-s))\pi$  for the limit of the conditional characteristic function. This is just one of the many ways that eventually lead to the conclusion that for  $\alpha \rightarrow \infty$  the process  $N^\alpha$  converge weakly to an ordinary Poisson process with constant intensity  $\lambda_\infty$ . In [24] one can find the stronger result that the variational distance between the MM law of  $N_t^\alpha, t \in [0, T]$  and the limit law is of order  $\alpha^{-1}$ .

## References

- [1] A. Ang and G. Bekaert (2002). Regime switches in interest rates. *Journal of Business and Economic Statistics* **20**, 163–182.
- [2] S. Asmussen (2003). *Applied Probability and Queues*, 2nd edition. Springer, New York.
- [3] S. Asmussen and H. Albrecher (2010). *Ruin probabilities*, 2nd edition. World Scientific, New Jersey.
- [4] T.R. Bielecki and M. Rutkowski (2002). *Credit risk: modelling, valuation and hedging*. Springer-Verlag, Berlin.
- [5] J. Blom, K. de Turck, and M. Mandjes (2015). Analysis of Markov-modulated infinite-server queues in the central-limit regime. *Probability in the Engineering and Informational Sciences* **29**, 433–459.
- [6] J. Blom, O. Kella, M. Mandjes, and H. Thorsdottir (2014). Markov-modulated infinite-server queues with general service times. *Queueing Systems* **76**, 403–424.
- [7] P. Brémaud (1980). *Point processes and Queues*. Springer-Verlag, New York-Berlin.
- [8] A. Chen and L. Delong (2015). Optimal Investment for a Defined-Contribution Pension Scheme under a Regime Switching Model, *ASTIN Bulletin*, March 2015, 1–23.
- [9] D. Duffie, D. Filipović, and W. Schachermayer (2003). Affine processes and applications in finance. *Ann. Appl. Probab.* **13**, 984–1053.
- [10] R.J. Elliott, L. Aggoun and J.B. Moore (1995). *Hidden Markov Models*. Springer-Verlag, New York.
- [11] R. Elliott and R. Mamon (2002). An interest rate model with a Markovian mean-reverting level. *Quantitative Finance* **2**, 454–458.
- [12] R. Elliott and T.K. Siu (2009). On Markov-modulated exponential-affine bond price formulae. *Applied Mathematical Finance* **16**, 1–15.
- [13] D. Filipović (2009). *Term-structure models*. Springer-Verlag, Berlin.
- [14] A. L. Goel and K. Okumoto (1978). Bayesian software prediction models, vol. 1: An imperfect debugging model for reliability and other quantitative measures of software systems. Rome Air Develop. Center, NY, *RADC-TR-78-155*.
- [15] G. Di Graziano (2014). Optimal Trading Stops and Algorithmic Trading. SSRN <http://ssrn.com/abstract=2381830>.
- [16] J. Hamilton (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* **57**, 357–384.
- [17] N.J. Higham (2008). *Functions of Matrices: Theory and Computation*. SIAM, Philadelphia.

- [18] G. Huang, M. Mandjes and P. Spreij (2014). Weak convergence of Markov-modulated diffusion processes with rapid switching. *Statist. Probab. Lett.* **86**, 74–79.
- [19] Gang Huang, Marijn Jansen, Michel Mandjes, Peter Spreij, Koen De Turck (2015). Markov-modulated Ornstein-Uhlenbeck processes. *Advances in Applied Probability* **48**(1), 235–254.
- [20] J. Jacod and A.N. Shiryaev (2003). *Limit theorems for stochastic processes, 2nd edition*. Grundlehren der Mathematischen Wissenschaften 288, Springer-Verlag, Berlin.
- [21] Z. Jelinski and P. B. Moranda (1972). Software reliability research. In *Statistical Computer Performance Evaluation*, W. Freiberger, Ed. NewYork: Academic, 465–484.
- [22] Z. Jiang, M.R. Pistorius (2008). On perpetual American put valuation and first passage in a regime-switching model with jumps. *Finance and Stochastics* **12**, 331–355.
- [23] Z. Jiang, M. Pistorius (2012). Optimal dividend distribution under Markov regime switching. *Finance and Stochastics* **16**, 449–476.
- [24] Y.M. Kabanov, R.S. Liptser, and A.N. Shiryaev (1983). Weak and strong convergence of distributions of counting processes. *Teor. Veroyatnost. i Primenen.*, **28**, 288–319.
- [25] G. Koch, P.J.C. Spreij (1983). Software Reliability as an Application of Martingale & Filtering Theory. *IEEE Transactions on Reliability* **R-32**, 342–345.
- [26] B. Littlewood (1980). Theories of Software Reliability: How Good Are They and How Can They Be Improved? *IEEE Transactions on Software Engineering* **SE-6**, 489–500.
- [27] M. Neuts (1981). *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. Courier Dover Publications, New York.
- [28] J.H. van Schuppen (1977). Filtering, prediction and smoothing for counting process observations, a martingale approach. *SIAM J. Appl. Math.* **32**, 552–570.
- [29] G. Yin and C. Zhu (2010). Hybrid Switching Diffusions: Properties and Applications. Springer, New York.
- [30] G. Yin (2009). Asymptotic expansions of option price under regime-switching diffusions with a fast-varying switching process, *Asymptotic Anal.* **65**, 203–222.
- [31] X.Y. Zhou and G. Yin (2003). Markowitz mean-variance portfolio selection with regime switching: A continuous-time model, *SIAM J. Control Optim.* **42**, 1466–1482.